Physically Inspired Gaussian Splatting for HDR Novel View Synthesis

Huimin Zeng[1] Yue Bai[1] Hailing Wang[1] Yun Fu[1,2]
[1] Department of Electrical and Computer Engineering, Northeastern University
[2]Khoury College of Computer Science, Northeastern University
{zeng.huim, bai.yue, wang.haili, y.fu}@northeastern.edu

Abstract

*High dynamic range novel view synthesis (HDR-NVS) reconstructs scenes with dynamic details by fusing multi-exposure low dynamic range (LDR) views, yet it struggles to capture ambient illumination-dependent appearance. Implicitly supervising HDR content by constraining tone-mapped results fails in correcting abnormal HDR values, and results in limited gradients for Gaussians in under/over-exposed regions. To this end, we introduce PhysHDR-GS, a physically inspired HDR-NVS framework that models scene appearance via intrinsic reflectance and adjustable ambient illumination. PhysHDR-GS employs a complementary image-exposure (IE) branch and Gaussian-illumination (GI) branch to faithfully reproduce standard camera observations and capture illumination-dependent appearance changes, respectively. During training, the proposed cross-branch HDR consistency loss provides explicit supervision for HDR content, while an illumination-guided gradient scaling strategy mitigates exposure-biased gradient starvation and reduces under-densified representations. Experimental results across realistic and synthetic datasets demonstrate our superiority in reconstructing HDR details (*e.g.*, a PSNR gain of **2.04 dB** over HDR-GS), while maintaining real-time rendering speed (up to **76 FPS**).*

1. Introduction

Novel view synthesis (NVS) [17, 25, 27, 31, 38] reconstructs a 3D scene from sparse image sequences and renders high-quality novel views, demonstrating wide applications in scenarios such as autonomous driving [1, 26] and AR/VR [9, 29]. However, due to the limited dynamic range of standard sensors, captured sequences are unable to faithfully reflect varying illumination in real-world scenes, resulting in limited details in highlights and shadows of the reconstructed scene. High dynamic range novel view synthesis (HDR-NVS) [3, 12–15, 23, 23, 28, 33, 37] addresses this by leveraging multi-exposures fusion (MEF) to combine complementary information from low dynamic range (LDR) images captured with different exposures. Recent advances from Neural Radiance Field (NeRF) [27] to 3D Gaussian Splatting (3DGS) [17] largely accelerate HDR-NVS task, enabling high-quality and real-time rendering.

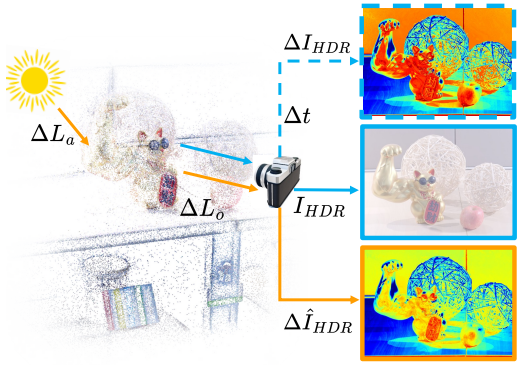


Figure 1. Variation of camera exposure Δt and ambient illumination ΔL_a scales the HDR signal in different ways: Δt causes a global change ΔI_{HDR}, while ΔL_a induces local changes $\Delta \hat{I}_{HDR}$ (*e.g.*, nameplate of the luckycat) via lighting-conditioned radiance variation ΔL_o. Their different response patterns reveal complementary ways of modeling dynamic-range details.

Despite high dynamic details introduced by MEF, HDR-NVS still faces several challenges for realistic reconstruction: *(1) appearance entanglement:* the appearance of objects is jointly determined by material properties and environmental conditions (*e.g.*, direct/indirect illumination). Simply scaling sensor shutter time (*i.e.*, exposure t) cannot disentangle these factors and reflect illumination-dependent appearance changes. As shown in Fig. 1, a change in exposure Δt mainly causes global intensity change ΔI_{HDR}, while ambient illumination change ΔL_a induces local, lighting-conditioned change $\Delta \hat{I}_{HDR}$ (*e.g.*, on the luckycat nameplate) via radiance variation ΔL_o; *(2) implicit HDR supervision:* HDR ground truth (GT) is typically unavailable. Supervision of reconstructed HDR view is therefore implicitly conducted by constraining LDR views tonemapped from HDR views. Since tone mapping compresses dynamic ranges, abnormal or saturated HDR values cannot be reflected in tonemapped results and be effectively

constrained; *(3) exposure-biased gradient starvation:* tone mapping curve typically yields small slopes at the extremes. Therefore, Gaussians covering under/over-exposed regions accumulate much smaller gradients than those at normal exposures (see Sec. 4.3). As a result, they struggle to meet densification thresholds, leading to under-densified representations and suboptimal reconstruction performance.

To address these challenges, we propose PhysHDR-GS, a physically inspired HDR-NVS framework that models the scene appearance as intrinsic reflectance and adjustable ambient illumination. Specifically, motivated by the different response patterns in Fig. 1, our framework contains two complementary branches: an image-exposure (IE) branch that modulates exposure on captured images, and a Gaussian-illumination (GI) branch that modulates the ambient illumination of 3D Gaussians. This dual modulation maintains fidelity to standard camera observations, while explicitly capturing ambient illumination-dependent appearance changes. During training, to provide explicit supervision for HDR content without GT, we impose a cross-branch HDR consistency loss between the pre-tonemapped HDR outputs of IE and GI branch, and further leverage a learnable tone mapper to fuse complementary details of tonemapped results as the LDR outputs. Moreover, to mitigate the exposure-biased gradient starvation, we introduce an illumination-guided gradient scaling strategy that amplifies per-Gaussian gradients based on illumination deviation (*i.e.*, discrepancy of intrinsic and relit illumination), preventing under/over-exposed Gaussians from being under-densified. Our contributions are summarized as follows:

- We propose a physically inspired HDR-NVS framework that combines an image-exposure (IE) branch and a Gaussian-illumination (GI) branch to reconstruct standard camera observations and explicitly capture illumination-dependent appearance changes, respectively.
- An HDR consistency loss is imposed between IE and GI branches to enable explicit HDR supervision without ground truth. A cross-fusion-based tone mapper further fuses LDR results to improve reconstruction quality.
- We propose an illumination-guided gradient scaling strategy that amplifies Gaussian gradients based on illumination deviation, alleviating gradient starvation and reducing under-densified representations.
- Experimental results on two exposure settings across three benchmarks demonstrate our superiority in reconstructing HDR details, while maintaining real-time rendering speed (up to **76 FPS**). Ablation studies further demonstrate the effectiveness of each component.

2. Related Work

2.1. High Dynamic Range Novel View Synthesis

Early NeRF-based HDR-NVS methods [12, 13, 15, 35] reconstruct an HDR radiance field from multi-exposure LDR inputs, with some exploring RAW space [28] to better preserve dynamic range. However, volumetric rendering makes both training and inference time-consuming. Recent 3DGS-based methods [14, 23, 23, 33] model scenes with Gaussian primitives and adopt rasterization-based rendering, achieving significant acceleration. HDR-GS [3] fits HDR color with spherical harmonics and predicts exposure-conditioned LDR views via an MLP-based tone mapper. Wu *et al*. [37] include luminance dimension for irradiance-to-color conversion and employ an asymmetric grid for tone mapping. To stabilize 3D tone mapping, GaussHDR [24] unifies 3D and 2D local tone mapping and fuses dual-branch LDR outputs. Currently, HDR-NVS has also been extended to multi-modal settings [5, 21], single-exposure [22, 41] and inconsistent illumination scenarios [2, 6, 11]. However, most existing methods still follow a conventional HDR imaging pipeline, where different lighting levels are simulated by applying exposure and tone mapping to 2D images. Without modeling illumination in 3D space, environment-dependent attributes of the scene are largely underexplored. Therefore, we explicitly model lighting-conditioned scene appearance by jointly controlling camera exposure and ambient illumination, revealing complementary HDR details in image space and 3D radiance.

2.2. High Dynamic Range Reconstruction

HDRCNN [8] makes an early exploration on reconstructing HDR from a single LDR image with a CNN. HDRNet [10] achieves real-time enhancement by learning a low-resolution-based affine color transform and applying it to the full-resolution image. For multi-exposure HDR imaging, recent methods [4, 7, 20] typically decompose this problem into alignment, fusion, and reconstruction. AFUNet [20] introduces a cross-iterative network that alternates alignment and fusion to progressively reconcile motion and exposure discrepancies. Debevec *et al*. [7] estimate the camera response curve from bracketed exposures under reciprocity and then fuse multiple photographs. Le *et al*. [19] invert the camera response and synthesize multi-exposure images to hallucinate missing details. SAFNet [18] selectively estimates cross-exposure motion and valuable region masks with shared decoders and performs explicit fusion. Generative approaches [30, 34, 36, 39] further leverage powerful priors to recover details in saturated or underexposed regions. Despite impressive HDR reconstruction demonstrated by these methods, they are essentially 2D methods and are unable to understand 3D scene, thus cannot synthesize novel HDR views.

3. Preliminaries

3.1. 3D Gaussian Splatting

3DGS [17] model a static 3D scene with a set of explicit Gaussian primitives G^{3D} as follows:

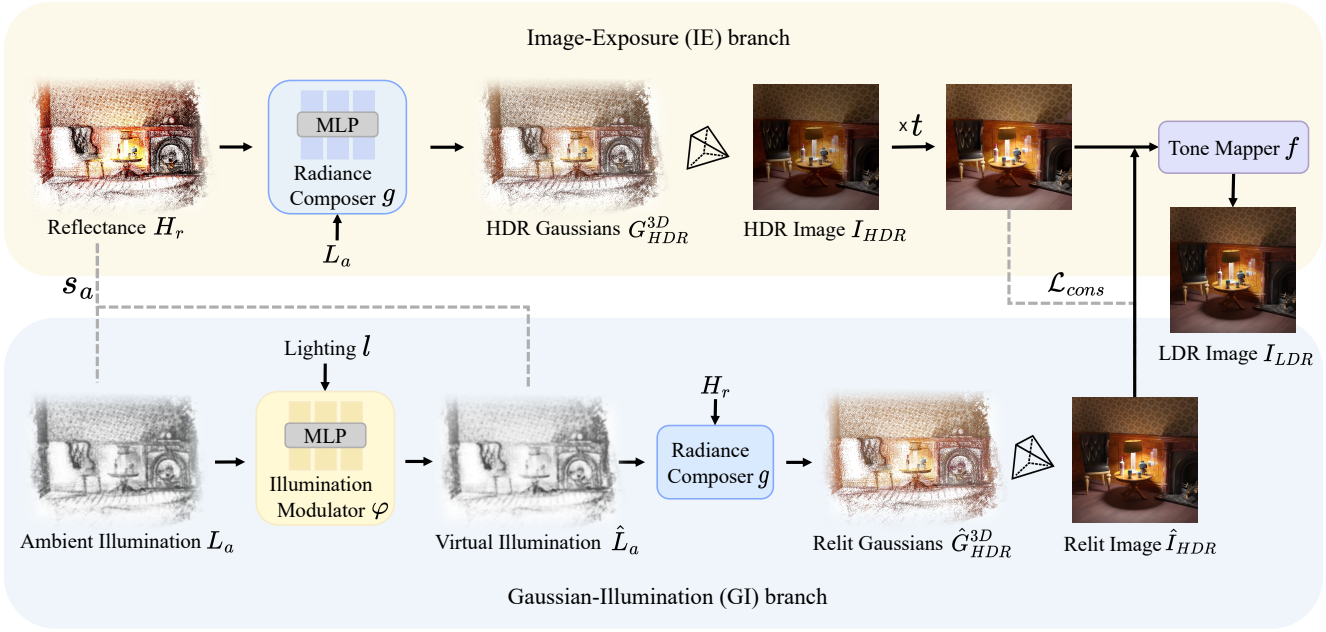


Figure 2. Overview of the proposed PhysHDR-GS, where Gaussian color is modeled from intrinsic reflectance and ambient illumination. The image–exposure (IE) branch modulates exposure t on 2D images, while the Gaussian–illumination (GI) branch modulates ambient illumination L_a on 3D Gaussians, yielding complementary dynamic-range details. Tone mapper f performs tone mapping and dual-branch fusion for final LDR results. During training, a cross-branch HDR consistency loss $\mathcal{L}_{cons}$ enables explicit HDR self-supervision. Illumination-Guided Gradient Scaling rescales per-Gaussian gradients with s_a to mitigate under-splitting in extreme exposure regions.

$$G^{3D} = \{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \boldsymbol{\alpha}_i, \boldsymbol{c}_i)\}_{i=1}^{N}, \tag{1}$$

where N is the number of Gaussians. $\boldsymbol{\mu}_i$, $\boldsymbol{\Sigma}_i$, $\boldsymbol{\alpha}_i$, $\boldsymbol{c}_i$ denote the center position, covariance, opacity and color of i-th Gaussian primitive, respectively. $\boldsymbol{\Sigma}_i$ is defined with a scaling matrix $\boldsymbol{S}_i$ and rotation matrix $\boldsymbol{R}_i$ as $\boldsymbol{R}_i\boldsymbol{S}_i\boldsymbol{S}_i^\top\boldsymbol{R}_i^\top$. During rendering, given the projection matrix $\boldsymbol{P}$, extrinsic matrix $\boldsymbol{W}$ and Jacobian of projective transformation $\boldsymbol{J}$, each 3D Gaussian primitive in world space is projected as a 2D Gaussian G^{2D}, with transformed covariance $\hat{\boldsymbol{\Sigma}} = \boldsymbol{JW\Sigma W}^\top\boldsymbol{J}^\top$ and camera coordinate $\hat{\boldsymbol{\mu}} = \boldsymbol{PW}[\boldsymbol{\mu}, 1]^\top \in \mathbb{R}^4$. Then the 2D pixel $\mathbf{p}$ is shaded as follows:

$$\boldsymbol{C}(\mathbf{p}) = \sum_{i=1}^{N} \boldsymbol{\alpha}_i \boldsymbol{c}_i G_i^{2D}(\mathbf{p}) \prod_{j=1}^{i-1} \left(1 - \boldsymbol{\alpha}_j G_j^{2D}(\mathbf{p})\right). \tag{2}$$

During training, the i-th primitive is densified if average screen-space gradient exceeds a threshold τ_p:

$$\frac{1}{M_i}\sum_{k=1}^{M_i}\left\|\frac{\partial \mathcal{L}_k}{\partial \boldsymbol{\mu}_{i,k}^{\text{ndc}}}\right\|_2 > \tau_p, \tag{3}$$

where M_i is the number of visible views, and $\mathcal{L}_k$ denote per-view loss. $\boldsymbol{\mu}^{\text{ndc}} = \frac{\hat{\boldsymbol{\mu}}_{1:3}}{\hat{\mu}_4}$ is the normalized device coordinates (NDC) center.

3.2. Physically-Based Rendering

Physically based rendering (PBR) models object color as the interaction between illumination and surface material. For a surface point $\mathbf{x}$ with normal $\mathbf{n}$, outgoing radiance toward direction $\boldsymbol{\omega}_o$ is given by rendering equation [16]:

$$L_o(\mathbf{x}, \boldsymbol{\omega}_o) = L_e(\mathbf{x}, \boldsymbol{\omega}_o) + \int_\Omega f_r(\mathbf{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o)\, L_i(\mathbf{x}, \boldsymbol{\omega}_i)\, (\mathbf{n}\boldsymbol{\omega}_i)\, d\boldsymbol{\omega}_i, \tag{4}$$

where L_e is emitted radiance, L_i is incident radiance from direction $\boldsymbol{\omega}_i$, f_r is the bidirectional reflectance distribution function (BRDF). Ω is the upper hemisphere around $\mathbf{n}$. With uniform hemispherical illumination (*i.e.*, $L_i(\mathbf{x}, \boldsymbol{\omega}_i) \equiv L_a(\mathbf{x})$ for all $\boldsymbol{\omega}_i \in \Omega$) and direction-independent emission (*i.e.*, $L_e(\mathbf{x}, \boldsymbol{\omega}_o) \equiv L_e(\mathbf{x})$), Eq. (4) simplifies to:

$$L_o(\mathbf{x}, \boldsymbol{\omega}_o) = L_e(\mathbf{x}) + L_a(\mathbf{x})\, H_r(\mathbf{x}, \boldsymbol{\omega}_o), \tag{5}$$

where $H_r(\mathbf{x}, \boldsymbol{\omega}_o) = \int_\Omega f_r(\mathbf{x}, \boldsymbol{\omega}_i, \boldsymbol{\omega}_o)\, (\mathbf{n}\boldsymbol{\omega}_i)\, d\boldsymbol{\omega}_i$ is the hemispherical-directional reflectance. A LDR pixel at location $\mathbf{p}$ is captured by applying exposure t and camera response function (CRF) f to the accumulated radiance:

$$I_{LDR}(\mathbf{p}) = f\big(t\, L_o(\mathbf{x}, \boldsymbol{\omega}_o)\big) = f\big(t\, L_e(\mathbf{x}) + t\, L_a(\mathbf{x})\, H_r(\mathbf{x}, \boldsymbol{\omega}_o)\big), \tag{6}$$

where $L_e(\mathbf{x})$ and $H_r(\mathbf{x}, \boldsymbol{\omega}_o)$ are scene-intrinsic and exposure-invariant, while t and $L_a(\mathbf{x})$ scales signal before CRF f. We model $L_o(\mathbf{x}, \boldsymbol{\omega}_o)$ with learnable parameters g, LDR pixel $\mathbf{p}$ thus depends on t and L_a as follows:

$$I_{LDR}(\mathbf{p}; t, L_a) = f\big(t \cdot g\,(L_a(\mathbf{x}), H_r(\mathbf{x}, \boldsymbol{\omega}_o))\big), \tag{7}$$

where constant L_e is absorbed. Eq. (7) shows that exposure t and ambient illumination L_a play complementary roles in shaping the dynamic range of the pre-CRF signal.

4. Method

Given a set of multi-exposure LDR views, we aim to reconstruct a 3D scene with HDR details. As shown in Fig. 2, Gaussian color is factorized into intrinsic reflectance H_r and adjustable ambient illumination L_a. To capture complementary dynamic details from exposure and ambient illumination, we introduce an image-exposure (IE) branch that follows the camera pipeline to modulate exposure t on 2D images, and a Gaussian–illumination (GI) branch that modulates ambient illumination L_a on 3D Gaussians. We elaborate the method from three aspects: physical radiance composition, self-consistent HDR fusion, and illumination-guided gradient scaling.

4.1. Physical Radiance Composition

Image-exposure (IE) branch. As shown in Fig. 2, given separately modeled reflectance H_r and ambient illumination L_a, the Gaussian color $\boldsymbol{c}$ is produced by an MLP-based radiance composer g:

$$\boldsymbol{c} = g(L_a, H_r). \tag{8}$$

Collecting geometry $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, opacity $\boldsymbol{\alpha}$ and color $\boldsymbol{c}$ further yields the HDR Gaussians set:

$$G^{3D}_{HDR} = \{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \boldsymbol{\alpha}_i, \boldsymbol{c}_i)\}^N_{i=1}. \tag{9}$$

Given a target view, the HDR Gaussian set G^{3D}_{HDR} is projected to HDR image I_{HDR}. To cover different luminance bands and bring mid-tone regions into the camera's responsive range, the IE branch applies exposure t on I_{HDR} (*i.e.*, $I_{HDR} \times t$) to globally scale the pre-tone-mapping signal.

Gaussian-illumination (GI) branch. With the disentanglement of reflectance and illumination, the 3D scene can be relit by modulating L_a based a target lighting condition. We introduce an illumination modulator φ to produce virtual illumination $\hat{L}_a$ as follows:

$$\hat{L}_a = \varphi(L_a, l), \tag{10}$$

where l denotes the target lighting level. Replacing L_a with $\hat{L}_a$ in Eq. (8) gives relit color $\hat{\boldsymbol{c}}$ and the relit Gaussians:

$$\hat{G}^{3D}_{HDR} = \{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \boldsymbol{\alpha}_i, \hat{\boldsymbol{c}}_i)\}^N_{i=1}, \tag{11}$$

which is further projected to relight image $\hat{I}_{HDR}$ with the same viewpoint of I_{HDR}. By adjusting virtual illumination $\hat{L}_a$, the GI branch enables rescaling radiance intensity locally to avoid saturation. Together, IE and GI branches provide better coverage of a higher dynamic range.

4.2. Self-Consistent HDR Fusion

Tone-mapped LDR learning. As illustrated in Fig. 2, imposing global exposure t and local virtual ambient illumination $\hat{L}_a$ yields HDR signals $I_{HDR} \times t$ and $\hat{I}_{HDR}$. To aggregate complementary dynamic range details and supervise with standard camera-captured LDR images, both HDR signals are further mapped into the LDR domain. As illustrated in Fig. 3, the proposed tone mapper f consists of two lightweight MLPs f_{tm} and f_{mix}, which perform global-local tone mapping and LDR fusion, respectively. Given an HDR image, f_{tm} conducts tone mapping and outputs a pair of global-local LDR images. Applying f_{tm} to the exposure-scaled $I_{HDR} \times t$ and relit $\hat{I}_{HDR}$ thus yields intermediate LDR predictions $\{I^{glo}_{LDR}, I^{loc}_{LDR}, \hat{I}^{glo}_{LDR}, \hat{I}^{loc}_{LDR}\}$. The fusion MLP f_{mix} then conducts cross-fusion:

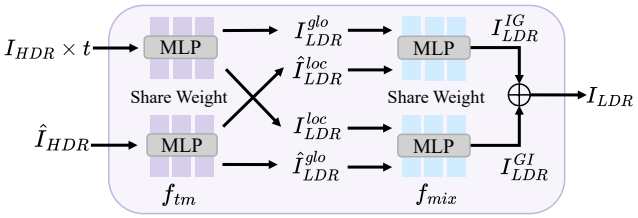


Figure 3. Illustration of the tone mapper f. Given inputs $I_{HDR} \times t$ and $\hat{I}_{HDR}$, the tone-mapping MLP f_{tm} first predicts global and local LDR outputs. The fusion MLP f_{mix} then cross-fuses these global-local pairs to produce the final LDR result I_{LDR}.

$$I^{IG}_{LDR} = f_{mix}(I^{glo}_{LDR}, \hat{I}^{loc}_{LDR}), \tag{12}$$

$$I^{GI}_{LDR} = f_{mix}(I^{glo}_{LDR}, I^{loc}_{LDR}). \tag{13}$$

The final LDR prediction I_{LDR} is obtained by aggregating the fused outputs:

$$I_{LDR} = I^{IG}_{LDR} + I^{GI}_{LDR}. \tag{14}$$

To supervise the results with available LDR views, we define the reconstruction loss as follows:

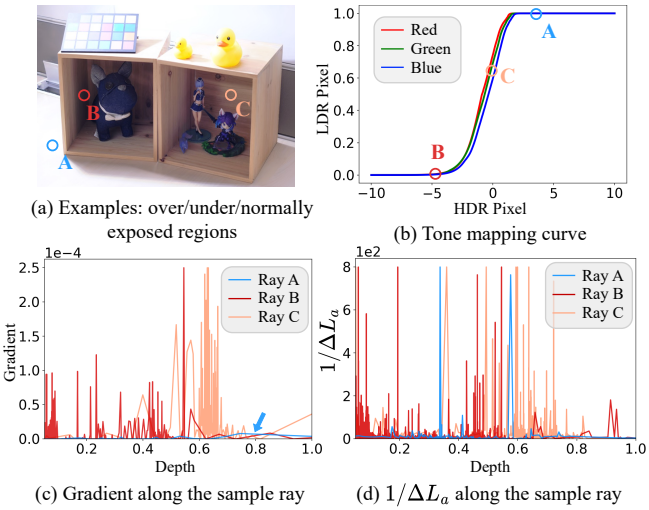
$$\mathcal{L}_{rec} = \sum_{I \in \mathcal{I}} \Big[\gamma\, \mathcal{L}_{MSE}(I, I_{gt}) + \mathcal{L}_{D\text{-}SSIM}(I, I_{gt})\Big], \tag{15}$$

where $\mathcal{I} = \{I_{LDR}, I^{IG}_{LDR}, I^{GI}_{LDR}\}$, I_{gt} is LDR ground truth. Hyperparameter γ controls the weight of MSE term.

Self-consistent HDR learning. Tone-mapping HDR images into LDR inevitably compresses dynamic range and clips extreme values (*e.g.*, saturated highlights and deep shadows). As a result, abnormal HDR values may not be effectively constrained by LDR supervision. To apply explicit supervision to HDR, we impose a cross-branch HDR consistency loss between the IE and GI branch. As shown in Fig. 2, for each view, we compute the consistent loss with exposure-scaled $I_{HDR} \times t$ and relit image $\hat{I}_{HDR}$ as follows,

$$\mathcal{L}_{\text{cons}} = \big\|\mathcal{G}\big(I_{HDR} \times t\big) - \mathcal{G}\big(\hat{I}_{HDR}\big)\big\|_1, \tag{16}$$

where $\mathcal{G}$ indicates applying Gaussian blur to avoid penalizing misaligned details. In practice, we set lighting level $l = t$ so that $I_{HDR} \times t$ and $\hat{I}_{HDR}$ are comparable in brightness. $\mathcal{L}_{\text{cons}}$ therefore matches the overall illumination and low-frequency structure of the HDR predictions, providing explicit supervision for the HDR content.



(a) Examples: over/under/normally exposed regions

(b) Tone mapping curve

(c) Gradient along the sample ray

(d) $1/\Delta L_a$ along the sample ray

Figure 4. Gradient and illumination deviation analysis, where over/under-exposed pixels lie in flat regions of the tone mapping curve and yield a small Gaussian gradient. The gradient shows positive correlation with reciprocal illumination deviation $1/\Delta L_a$.

4.3. Illumination-Guided Gradient Scaling

Given multi-exposure views of the same scene, pixels can be categorized into three types: over-exposed, under-exposed and normally-exposed regions (*e.g.*, point A, B, C in Fig. 4(a)). The tone mapping curve (shown in Fig. 4(b)) typically compresses dynamic range, clipping highlights and shadows. Consequently, compared with normally-exposed regions (*e.g.*, point C), over-exposed (*e.g.*, point B) and under-exposed regions (*e.g.*, point A) lie in flat region of CRF (see Fig. 4(b)) and yiled much smaller slopes. As shown in Fig. 4(c), Gaussian primitives that model these regions receive very limited gradient. Due to the gradient-based densification shown in Eq. (3), these Gaussians seldom reach the threshold τ_p, resulting in insufficient splitting and under-densified representations [32, 40, 43].

We observe that the illumination deviation of a Gaussian primitive is inversely correlated to the received gradient. As shown in Fig. 4(d), we visualize the reciprocal of illumination deviation $1/\Delta L_a$, where $\Delta L_a = |L_a - \hat{L}_a|$ measures the illumination discrepancy. Compared with normally exposed regions, over-exposed/under-exposed regions exhibit much larger illumination deviation (*i.e.*, smaller $1/\Delta L_a$), which is negatively related to their gradient magnitudes. This suggests an illumination deviation-based strategy to compensate for gradient deficiency. Based on this observation, we propose the illumination-guided gradient scaling, which adaptively rescales per-Gaussian gradients. The gradient scaling factor s_a is defined as:

$$s_a \;=\; s \cdot \sigma\big(|L_a - \hat{L}_a|\big) + 1, \tag{17}$$

where $\sigma(\cdot)$ is the sigmoid function, s is a hyperparameter controlling the maximum scaling strength. Therefore, the densification criterion in Eq. (3) is revised below:

$$\mathbb{I}_i\left(s_a\right) \frac{1}{M_i} \sum_{k=1}^{M_i} \left\| \frac{\partial \mathcal{L}_k}{\partial \boldsymbol{\mu}_{i,k}^{\text{ndc}}} \right\|_2 > \tau_p, \tag{18}$$

where $\mathbb{I}_i\left(s_a\right)$ indicates retrieving the scaling factor for i-th Gaussian primitive. By amplifying gradients according to the illumination deviation, illumination-guided gradient scaling effectively prevents Gaussians in over/under-exposed regions from insufficient splitting, leading to superior reconstruction performance.

4.4. Loss Function

For each training view, the total loss combines the LDR reconstruction loss and the HDR consistency loss. For synthetic datasets, we follow [3, 12, 24] to additionally impose a uniform-exposure regularization:

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_{\text{rec}} + \lambda_2 \mathcal{L}_{\text{cons}} + \lambda_3 \mathcal{L}_{\text{unit}}, \tag{19}$$

where λ_1, λ_2 and λ_3 are weight of each loss term .

5. Experiment

5.1. Experimental Settings

Datasets and settings. Experiments are conducted on three benchmarks: realistic HDR-NeRF-Real [12], HDR-Plenoxels-Real [15] and synthesized HDR-NeRF-Syn [12]. All datasets provide LDR views at five exposure times $\{t_1, t_2, t_3, t_4, t_5\}$, which are grouped into LDR-OE ($\{t_1, t_3, t_5\}$) and LDR-NE ($\{t_2, t_4\}$). HDR-NeRF-Syn additionally provides HDR ground truth for evaluating HDR content. Note that these HDR ground-truth images are only used for evaluation and are not used for training by any method. For HDR-NeRF-Real and HDR-NeRF-Syn dataset, we train with 18 LDR-OE views and evalute with remaining 17 views. For HDR-Plenoxels-Real dataset, we train with 27 LDR-OE views and evaluate with remaining 13 views. We follow GaussHDR [24] to train with two exposure settings: randomly assigning exposure at $\{t_1, t_3, t_5\}$ at every iteration, (denoted as exp3), and initializing with random exposure from $\{t_1, t_3, t_5\}$ and keeping fixed throughout training (denoted as exp1).

Baselines and model variants. We compared against both NeRF-based (*i.e.*, HDR-NeRF [12]) and 3DGS-based methods. For the 3DGS-based methods methods, we include HDR-GS [3] and GaussHDR [24]).To assess the generality, we follow GaussHDR to integrate proposed components into both vanilla 3DGS backbone and Scaffold-GS [25]. Variants built on Scaffold-GS are marked with †.

Metrics. We evaluate reconstruction quality using PSNR and SSIM for pixel-wise fidelity, and LPIPS [42] for perceptual quality. HDR views are first tone-mapped with μ-law and then conduct quantitative evaluation [12, 24].

Table 1. Quantitative results on realistic HDR-NeRF-Real [12] and HDR-Plenoxels-Real [15] datasets, where the best and second-best results are highlighted in red and yellow, respectively. † indicates variants built on Scaffold-GS [25]. The proposed method achieves the overall best performance, demonstrating its effectiveness in synthesizing high-quality novel views across different exposure levels.

	Method	HDR-NeRF-Real						HDR-Plenoxels-Real					
		LDR-OE (t_1, t_3, t_5)			LDR-NE (t_2, t_4)			LDR-OE (t_1, t_3, t_5)			LDR-NE (t_2, t_4)		
		PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
exp3	HDR-NeRF [12]	34.27	0.9532	0.063	32.15	0.9475	0.074	-	-	-	-	-	-
	HDR-GS [3]	34.87	0.9697	0.021	31.02	0.9636	0.029	31.17	0.9509	0.040	28.60	0.9285	0.051
	GaussHDR [24]	36.05	0.9739	0.015	33.49	0.9707	0.017	31.50	0.9530	0.037	28.92	0.9326	0.044
	GaussHDR† [24]	36.32	0.9770	0.011	33.84	0.9738	0.014	32.87	0.9590	0.028	29.78	0.9391	0.037
	Ours	36.19	0.9779	0.011	33.68	0.9738	0.014	32.19	0.9536	0.033	28.79	0.9299	0.046
	Ours†	36.91	0.9777	0.009	34.15	0.9737	0.012	33.06	0.9592	0.025	29.76	0.9389	0.034
exp1	HDR-NeRF [12]	34.26	0.9532	0.063	31.55	0.9483	0.074	-	-	-	-	-	-
	HDR-GS [3]	32.96	0.9597	0.028	29.66	0.9535	0.035	29.62	0.9373	0.051	27.32	0.9151	0.060
	GaussHDR [24]	34.59	0.9678	0.018	32.98	0.9653	0.020	30.87	0.9458	0.041	28.26	0.9230	0.050
	GaussHDR† [24]	34.63	0.9710	0.014	33.29	0.9684	0.016	32.24	0.9540	0.031	28.88	0.9315	0.041
	Ours	34.64	0.9707	0.015	32.64	0.9672	0.018	31.20	0.9464	0.037	27.59	0.9193	0.047
	Ours†	34.84	0.9705	0.012	33.40	0.9677	0.014	32.34	0.9540	0.029	29.06	0.9313	0.038

Table 2. Quantitative results on the synthetic HDR-NeRF-Syn [12] dataset, where the best and second-best results are highlighted in red and yellow, respectively. † indicates variants built on Scaffold-GS [25]. Our method consistently outperforms baselines on both LDR and HDR scenarios, demonstrating its effectiveness in reconstructing HDR details and preserving information during tonemapping.

	Method	LDR-OE (t_1, t_3, t_5)			LDR-NE (t_2, t_4)			HDR		
		PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
exp3	HDR-NeRF [12]	38.82	0.9657	0.032	38.07	0.9641	0.034	26.63	0.9523	0.046
	HDR-GS [3]	40.28	0.9781	0.018	27.07	0.8744	0.127	17.51	0.6982	0.205
	GaussHDR [24]	42.28	0.9853	0.007	41.65	0.9850	0.007	37.78	0.9704	0.017
	GaussHDR† [24]	43.87	0.9899	0.004	42.74	0.9894	0.004	39.08	0.9767	0.011
	Ours	43.11	0.9848	0.008	42.48	0.9852	0.008	38.77	0.9710	0.016
	Ours†	44.26	0.9899	0.003	43.19	0.9896	0.004	39.21	0.9768	0.010
exp1	HDR-NeRF [12]	38.68	0.9649	0.032	37.63	0.9621	0.035	26.61	0.9523	0.046
	HDR-GS [3]	38.22	0.9688	0.026	25.42	0.8639	0.136	16.43	0.6048	0.265
	GaussHDR [24]	41.67	0.9839	0.007	41.21	0.9840	0.008	37.44	0.9691	0.018
	GaussHDR† [24]	42.94	0.9883	0.004	42.01	0.9879	0.005	38.62	0.9752	0.011
	Ours	41.99	0.9835	0.006	41.47	0.9842	0.007	38.34	0.9701	0.013
	Ours†	42.80	0.9881	0.004	42.13	0.9882	0.004	38.79	0.9756	0.010

Implementation details. We train each scene for 30k iterations. During the first 10k iterations, we freeze the fusion MLP f_{mix} and only train the tone-mapping MLP f_{tm}. f_{mix} is then unfrozen for the remaining iterations. The weights λ_1, λ_2 and λ_3 are set to 1, 0.5 and 0, respectively. For synthetic scenes, we additionally enable the unit-exposure regularization term by setting λ_3 set to 0.5. γ is set to 0.2. Scaling hyperparameter s is set to 1. All experiments are conducted with a single A6000 GPU. More details are included in the supplementary materials.

5.2. Quantitative Results

Quantitative results on the realistic HDR-NeRF-Real and HDR-Plenoxels-Real datasets are included in Tab. 1. As can be seen, the proposed Scaffold-GS variant (*i.e.*, Ours†) achieves the overall best performance under both exposure settings, yielding a PSNR gain of 0.59 dB over GaussHDR† on LDR-OE/exp3 of HDR-NeRF-Real. The 3DGS variant (*i.e.*, Ours) also demonstrates competitive performance, showing even better performance than Scaffold-GS-based GaussHDR† on LDR-OE/exp1 of HDR-NeRF-Real. Results on synthetic HDR-NeRF-Syn are reported in Tab. 2, where Ours† shows consistent superiority over compared methods (*e.g.*, a PSNR gain of 0.45 dB on LDR-NE/exp3 over GaussHDR†), achieving the overall leading performance. It's worth noting that Ours† shows consistent superiority over all compared methods on the perceptual LPIPS across all benchmarks, while Ours also demonstrates competitive perceptual performance and surpasses GaussHDR on most metrics. This highlights the benefit of jointly modeling image exposure and ambient illumination for

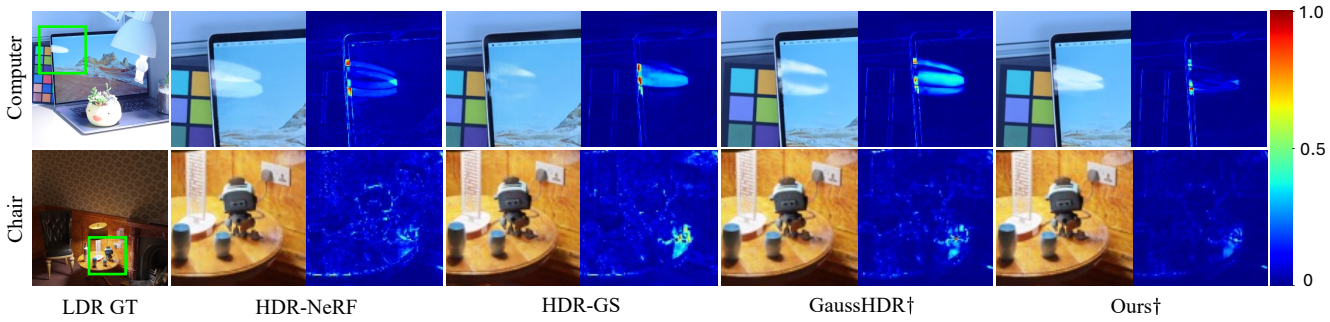


Figure 5. Qualitative comparisons on LDR views. For each method, we show the reconstructed LDR image and the residual map w.r.t. the ground truth. Competing methods exhibit noticeable missing content in saturated regions (e.g., screen reflections in the 1st row), indicating information loss after tone mapping, whereas our method effectively preserves fine structures and details.

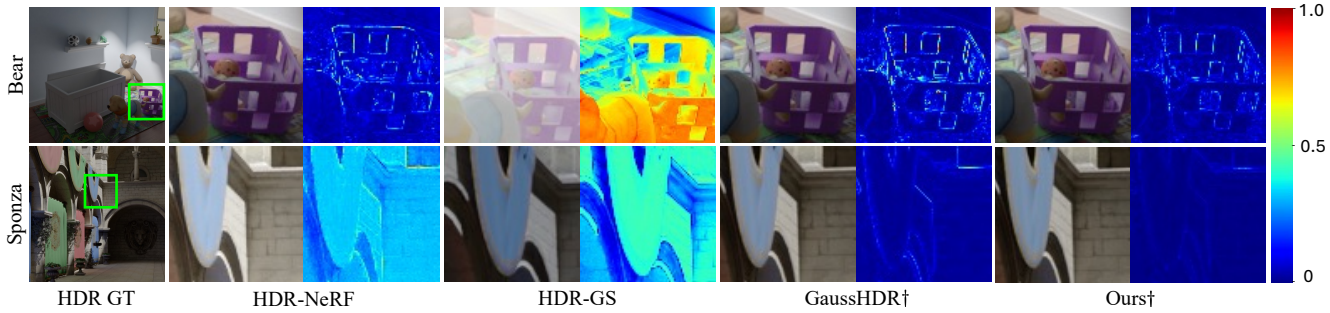


Figure 6. Qualitative comparisons on HDR views, where we include residual maps between the results and GT to highlight the difference. HDR-NeRF and HDR-GS struggle to reproduce correct illumination levels due to the absence of HDR supervision during training, leading to inaccurate brightness and lost details. By imposing cross-branch HDR consistency, our method faithfully estimates scene lighting and reconstructs fine structures (*e.g.*, basket edges in the 1st row).

HDR-NVS, enabling both faithful LDR reconstruction and perceptually pleasing dynamic-range detail, particularly in challenging highlight regions (see Sec. 5.3). Detailed quantitative results of each scene are reported in the supplementary materials.

5.3. Qualitative Results

Qualitative comparisons on LDR views are included in Fig. 5 and Fig. 6, respectively. For each method, we also visualize the residual map with respect to the GT to intuitively demonstrate the difference. As can be seen, for LDR scenes, the compared methods exhibit noticeable missing contents for saturated regions (*e.g.*, the reflection region in Fig. 5), indicating unconstrained information loss after tone mapping. In contrast, our method effectively preserves fine structures. We attribute this to the cross-branch HDR consistency loss that provides explicit supervision for regions of high signal intensity. For HDR scenes, methods such as HDR-NeRF and HDR-GS struggle to reproduce the correct illumination level, which results from the absence of HDR GT during training. While our method accurately estimates the lighting conditions by explicitly modeling exposure and ambient illumination, yielding effective reconstruction of details (*e.g.*, edge of basket in the 1st row) over GaussHDR†. More qualitative comparisons are included in the supplementary materials.

Table 3. Efficiency analysis measured on a single NVIDIA A6000 GPU. Rendering time and throughput are evaluated at an output resolution of 400×400.

Method	Rendering (ms)	Throughput (FPS)	Training (min)	Memory (MB)
HDR-NeRF	4189	0.24	500	11049
HDR-GS	9	117	10	5014
GaussHDR	19	53	28	5596
GaussHDR†	26	38	21	6221
Ours	13	76	15	3274
Ours†	19	53	18	3920

5.4. Efficiency Analysis

We conduct a comprehensive efficiency comparison in Tab. 3, reporting rendering speed, throughput, training time and GPU memory usage. Despite the additional radiance composition step, our method remains efficient. Ours and Ours† achieve real-time throughput of 76 FPS and 53 FPS at a target resolution of 400×400, which is $322\times$ and $220\times$ faster than HDR-NeRF, respectively. Compared with efficient HDR-GS, our method provides promising perfor-

Table 4. Ablation studies on HDR-NeRF-Real and HDR-Plenoxels-Real dataset, with the best and second-best results highlighted in red and yellow. IE branch indicates the baseline containing only the image-exposure branch. GI branch, HDR-cons and I-GS denote Gaussian-illumination branch, self-consistent HDR learning and illumination-guided gradient scaling, respectively.

Method	HDR-NeRF-Real						HDR-Plenoxels-Real					
	LDR-OE (t_1, t_3, t_5)			LDR-NE (t_2, t_4)			LDR-OE (t_1, t_3, t_5)			LDR-NE (t_2, t_4)		
	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓
IE branch	36.18	0.9766	0.0105	33.38	0.9727	0.0142	32.73	0.9576	0.0274	29.47	0.9368	0.0358
+ GI branch	36.27	0.9770	0.0096	33.46	0.9728	0.0132	32.83	0.9570	0.0271	29.34	0.9350	0.0362
+ HDR-cons	36.43	0.9774	0.0094	33.84	0.9732	0.0129	32.97	0.9584	0.0262	29.45	0.9364	0.0354
+ I-GS	36.91	0.9777	0.0093	34.15	0.9737	0.0124	33.06	0.9592	0.0253	29.76	0.9389	0.0339

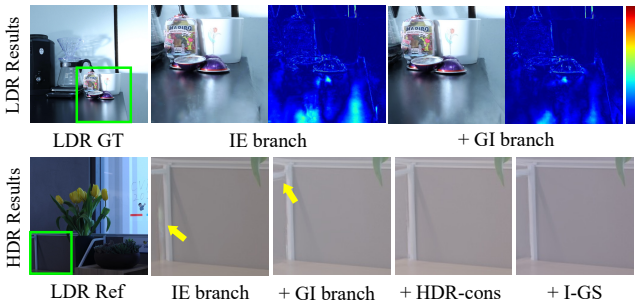


Figure 7. Visualization of ablation studies, where residual maps between LDR results and LDR GT are provided. Including GI branch effectively captures lighting-dependent appearance (*e.g.*, table reflections in 1st row) and reduces color distortions. Introducing HDR-cons and I-GS further refines structural details.

mance improvement (*e.g.*, 3.98 dB PSNR gain on HDR-NeRF-Syn/exp3) with tolerable reduction in speed. Compared with well-performing GaussHDR, our models deliver faster rendering speed (*e.g.*, Ours is 1.43× faster than GaussHDR) and lower GPU memory consumption, while still offering overall better reconstruction performance.

5.5. Ablation Studies

We conduct the ablation studies starting from a baseline that contains the IE branch only (denoted as *IE branch*). Then we progressively equip this baseline with the GI branch, cross-branch HDR consistency loss, and illumination-guided gradient scaling to evaluate the contribution of each component. All ablations are conducted under the exposure setting of exp3 across HDR-NeRF-Real and HDR-Plenoxels-Real datasets for generality.

Gaussian-illumination branch. As shown in Tab. 4, introducing Gaussian-illumination branch (denoted as + *GI branch*) yields consistent gains on HDR-NeRF-Real and overall improvement on HDR-Plenoxels-Real. We attribute this to the enhanced ability of the full framework to capture both exposure-scaled and lighting-conditioned appearance variations. As shown Fig. 7, including the GI branch effectively improves the modeling of the lighting-dependent appearance (*e.g.*, table surface reflection in the 1st row) and corrects texture distortion (*e.g.*, desk shelf in the 2nd row).

Self-consistent HDR learning. We verify the effectiveness of self-consistent HDR learning by imposing the cross-branch HDR consistency loss $\mathcal{L}_{cons}$ between IE and GI branches (denoted as + *HDR-cons*). As shown in Tab. 4, imposing consistency loss leads to noticeable improvement, with a maximum PSNR gain of 0.38 dB on HDR-NeRF-Real. Qualitative results in Fig. 7 indicate that self-consistent HDR learning helps to refine texture distortions.

Illumination-guided gradient scaling. We further include illumination-guided gradient scaling (denoted as + *I-GS*) during training to verify its effectiveness. As can be seen As can be seen from Fig. 7, scaling the per-Gaussian gradient with the proposed I-GS strategy introduces significant improvement, with a maximum PSNR gain of 0.48 dB and 0.31 dB on HDR-NeRF-Real and HDR-Plenoxels-Real, yielding the best overall configuration. Visualization in Fig. 7 shows that I-GS effectively alleviates texture distortions in shadow regions by protecting Gaussians from insufficient splitting.

6. Conclusion

In this paper, we propose a physically inspired HDR-NVS framework. To reflect the illumination-dependent appearance, we decompose the Gaussian color into intrinsic reflectance and adjustable ambient illumination. An image-exposure branch and a Gaussian-illumination branch jointly model complementary dynamic range details. An HDR consistency loss is imposed across these branches to enable self-supervision for HDR contents. The proposed gradient scaling strategy further amplifies the Gaussian gradients to prevent under-densified representations. Extensive experiments show that our method effectively reconstructs dynamic range details, while keeping the real-time rendering speed. Ablation studies further verify the effectiveness of each proposed component in our framework.

References

[1] Mohammad Altillawi, Fengyi Shen, Liudi Yang, Sai Manoj Prakhya, and Ziyuan Liu. Ce-npbg: Connectivity enhanced neural point-based graphics for novel view synthesis in autonomous driving scenes. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 2551–2559, 2025. 1

[2] Christophe Bolduc, Yannick Hold-Geoffroy, and Jean-François Lalonde. Gaslight: Gaussian splats for spatially-varying lighting in hdr. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 29120–29130, 2025. 2

[3] Yuanhao Cai, Zihao Xiao, Yixun Liang, Minghan Qin, Yulun Zhang, Xiaokang Yang, Yaoyao Liu, and Alan L Yuille. Hdr-gs: Efficient high dynamic range novel view synthesis at 1000x speed via gaussian splatting. *Advances in Neural Information Processing Systems*, 37:68453–68471, 2024. 1, 2, 5, 6

[4] Rufeng Chen, Bolun Zheng, Hua Zhang, Quan Chen, Chenggang Yan, Gregory Slabaugh, and Shanxin Yuan. Improving dynamic hdr imaging with fusion transformer. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 340–349, 2023. 2

[5] Zehao Chen, Zhan Lu, De Ma, Huajin Tang, Xudong Jiang, Qian Zheng, and Gang Pan. Evhdr-gs: Event-guided hdr video reconstruction with 3d gaussian splatting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2367–2375, 2025. 2

[6] Ziteng Cui, Xuangeng Chu, and Tatsuya Harada. Luminance-gs: Adapting 3d gaussian splatting to challenging lighting conditions with view-adaptive curve adjustment. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 26472–26482, 2025. 2

[7] Paul E Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pages 643–652. 2023. 2

[8] Gabriel Eilertsen, Joel Kronander, Gyorgy Denes, Rafał K Mantiuk, and Jonas Unger. Hdr image reconstruction from a single exposure using deep cnns. *ACM transactions on graphics (TOG)*, 36(6):1–15, 2017. 2

[9] Yonggan Fu, Yuecheng Li, Chenghui Li, Jason Saragih, Peizhao Zhang, Xiaoliang Dai, and Yingyan Celine Lin. Auto-card: Efficient and robust codec avatar driving for real-time mobile telepresence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21036–21045, 2023. 1

[10] Michaël Gharbi, Jiawen Chen, Jonathan T Barron, Samuel W Hasinoff, and Frédo Durand. Deep bilateral learning for real-time image enhancement. *ACM Transactions on Graphics (TOG)*, 36(4):1–12, 2017. 2

[11] Shucheng Gong, Lingzhe Zhao, Wenpu Li, Hong Xie, Yin Zhang, Shiyu Zhao, and Peidong Liu. Casualhdrsplat: Robust high dynamic range 3d gaussian splatting from casually captured videos. *arXiv preprint arXiv:2504.17728*, 2025. 2

[12] Xin Huang, Qi Zhang, Ying Feng, Hongdong Li, Xuan Wang, and Qing Wang. Hdr-nerf: High dynamic range neural radiance fields. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18398–18408, 2022. 1, 2, 5, 6

[13] Xin Huang, Qi Zhang, Ying Feng, Hongdong Li, and Qing Wang. Ltm-nerf: Embedding 3d local tone mapping in hdr neural radiance field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):10944–10959, 2024. 2

[14] Xin Jin, Pengyi Jiao, Zheng-Peng Duan, Xingchao Yang, Chongyi Li, Chun-Le Guo, and Bo Ren. Lighting every darkness with 3dgs: Fast training and real-time rendering for hdr view synthesis. *Advances in Neural Information Processing Systems*, 37:80191–80219, 2024. 2

[15] Kim Jun-Seong, Kim Yu-Ji, Moon Ye-Bin, and Tae-Hyun Oh. Hdr-plenoxels: Self-calibrating high dynamic range radiance fields. In *European Conference on Computer Vision*, pages 384–401. Springer, 2022. 1, 2, 5, 6

[16] James T Kajiya. The rendering equation. In *Proceedings of the 13th annual conference on Computer graphics and interactive techniques*, pages 143–150, 1986. 3

[17] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):139–1, 2023. 1, 2

[18] Lingtong Kong, Bo Li, Yike Xiong, Hao Zhang, Hong Gu, and Jinwei Chen. Safnet: Selective alignment fusion network for efficient hdr imaging. In *European Conference on Computer Vision*, pages 256–273. Springer, 2024. 2

[19] Phuoc-Hieu Le, Quynh Le, Rang Nguyen, and Binh-Son Hua. Single-image hdr reconstruction by multi-exposure generation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 4063–4072, 2023. 2

[20] Xinyue Li, Zhangkai Ni, and Wenhan Yang. Afunet: Cross-iterative alignment-fusion synergy for hdr reconstruction via deep unfolding paradigm. *arXiv preprint arXiv:2506.23537*, 2025. 2

[21] Yunhao Li, Xiang Liu, Xiaodong Wang, Xin Yuan, and Peidong Liu. Learning radiance fields from a single snapshot compressive image. *arXiv preprint arXiv:2412.19483*, 2024. 2

[22] Yiyu Li, Haoyuan Wang, Ke Xu, Gerhard Petrus Hancke, and Rynson WH Lau. Sehdr: Single-exposure hdr novel view synthesis via 3d gaussian bracketing. *arXiv preprint arXiv:2509.20400*, 2025. 2

[23] Zhihao Li, Yufei Wang, Alex Kot, and Bihan Wen. From chaos to clarity: 3dgs in the dark. *Advances in Neural Information Processing Systems*, 37:94971–94992, 2024. 1, 2

[24] Jinfeng Liu, Lingtong Kong, Bo Li, and Dan Xu. Gausshdr: High dynamic range gaussian splatting via learning unified 3d and 2d local tone mapping. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 5991–6000, 2025. 2, 5, 6

[25] Tao Lu, Mulin Yu, Linning Xu, Yuanbo Xiangli, Limin Wang, Dahua Lin, and Bo Dai. Scaffold-gs: Structured 3d gaussians for view-adaptive rendering. In *Proceedings of*

the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 20654–20664, 2024. 1, 5, 6

[26] Xin Ma, Jiguang Zhang, Peng Lu, Shibiao Xu, and Chengwei Pan. Novel view synthesis under large-deviation viewpoint for autonomous driving. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 6000–6008, 2025. 1

[27] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106, 2021. 1

[28] Ben Mildenhall, Peter Hedman, Ricardo Martin-Brualla, Pratul P Srinivasan, and Jonathan T Barron. Nerf in the dark: High dynamic range view synthesis from noisy raw images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 16190–16199, 2022. 1, 2

[29] Norman Müller, Katja Schwarz, Barbara Rössle, Lorenzo Porzi, Samuel Rota Bulo, Matthias Nießner, and Peter Kontschieder. Multidiff: Consistent novel view synthesis from a single image. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10258–10268, 2024. 1

[30] Yuzhen Niu, Jianbin Wu, Wenxi Liu, Wenzhong Guo, and Rynson WH Lau. Hdr-gan: Hdr image reconstruction from multi-exposed ldr images with large motions. *IEEE Transactions on Image Processing*, 30:3885–3896, 2021. 2

[31] Avinash Paliwal, Wei Ye, Jinhui Xiong, Dmytro Kotovenko, Rakesh Ranjan, Vikas Chandra, and Nima Khademi Kalantari. Coherentgs: Sparse novel view synthesis with coherent 3d gaussians. In *European Conference on Computer Vision*, pages 19–37. Springer, 2024. 1

[32] Julien Philip and Valentin Deschaintre. Floaters no more: Radiance field gradient scaling for improved near-camera training. *arXiv preprint arXiv:2305.02756*, 2023. 5

[33] Shreyas Singh, Aryan Garg, and Kaushik Mitra. Hdr-splat: Gaussian splatting for high dynamic range 3d scene reconstruction from raw images. *arXiv preprint arXiv:2407.16503*, 2024. 1, 2

[34] Chao Wang, Ana Serrano, Xingang Pan, Bin Chen, Karol Myszkowski, Hans-Peter Seidel, Christian Theobalt, and Thomas Leimkühler. Glowgan: Unsupervised learning of hdr images from ldr images in the wild. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10509–10519, 2023. 2

[35] Chao Wang, Krzysztof Wolski, Bernhard Kerbl, Ana Serrano, Mojtaba Bemana, Hans-Peter Seidel, Karol Myszkowski, and Thomas Leimkühler. Cinematic gaussians: Real-time hdr radiance fields with depth of field. In *Computer Graphics Forum*, page e15214. Wiley Online Library, 2024. 2

[36] Chao Wang, Zhihao Xia, Thomas Leimkuhler, Karol Myszkowski, and Xuaner Zhang. Lediff: Latent exposure diffusion for hdr generation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 453–464, 2025. 2

[37] Jiahao Wu, Lu Xiao, Rui Peng, Kaiqiang Xiong, and Ronggang Wang. Hdrgs: High dynamic range gaussian splatting. *arXiv preprint arXiv:2408.06543*, 2024. 1, 2

[38] Jiacong Xu, Yiqun Mei, and Vishal Patel. Wild-gs: Real-time novel view synthesis from unconstrained photo collections. *Advances in Neural Information Processing Systems*, 37:103334–103355, 2024. 1

[39] Huimin Zeng, Xinliang Zhang, Zhibin Yu, and Yubo Wang. Sr-itm-gan: Learning 4k uhd hdr with a generative adversarial network. *IEEE Access*, 8:182815–182827, 2020. 2

[40] Huimin Zeng, Yue Bai, and Yun Fu. Arbitrary-scale 3d gaussian super-resolution. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 12304–12312, 2026. 5

[41] Kaixuan Zhang, Hu Wang, Minxian Li, Mingwu Ren, Mao Ye, and Xiatian Zhu. High dynamic range novel view synthesis with single exposure. *arXiv preprint arXiv:2505.01212*, 2025. 2

[42] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018. 5

[43] Zheng Zhang, Wenbo Hu, Yixing Lao, Tong He, and Hengshuang Zhao. Pixel-gs: Density control with pixel-aware gradient for 3d gaussian splatting. In *European Conference on Computer Vision*, pages 326–342. Springer, 2024. 5